Un (no tan corto) resumen de Análisis I

Julián Sackmann

10/04/2012

1 Introducción

Proposición 1 (Principio de Arquimedianeidad)**.** *Si x es un número real, entonces existe $n \in \mathbb{N}$ tal que $n > x$.*

Demostración. Sean $x \in \mathbb{R}$ y $X = \{k \in \mathbb{N} : k \leq x\}$. Si $X = \emptyset$, entonces el principio vale pues $1 > x$ y se puede considerar $n = 1$. Si $x \neq \emptyset$, como está acotado superiormente (por x), tiene un supremo $s = \sup X \in \mathbb{R}$. Necesariamente debe existir $k_0 \in X \in \mathbb{N}$ tal que $s - 1 < k_0$ (pues $s - 1$ no es cota superior). Tomando $n = k_0 + 1$ se tiene $n \in \mathbb{N}$ y $n > s$. O sea $n \notin X$ lo que implica que $n > x$. $\square$

Proposición 2 (Equivalencia de Supremo)**.** *s es el supremo de A si y sólo si:*

- *$s \geq a$ para todo $a \in A$*
- *dado $\varepsilon > 0$ existe $a \in A$ tal que $s - \varepsilon < a$*

Demostración.
$\Rightarrow$) El primer ítem es trivial. Por otro lado, dado $\varepsilon > 0$, se tienen $s - \varepsilon < s$, por lo que $s - e$ no es cota superior de A. O sea, existe $a \in A$ tal que $s - \varepsilon < a$, lo que prueba el segundo ítem.

$\Leftarrow$) Es necesario probar que s es la cota superior más chica. Sea s' una cota superior más chica que s (O sea $s' < s$) y sea $\varepsilon = s - s' > 0$. Esto es absurdo porque contradice la hipótesis de que para cualquier $\varepsilon > 0$, existe $a \in A$ tal que $s - \varepsilon < a$. $\square$

2 Sucesiones

Proposición 3. *Si $s = sup\ A$ entonces existe una sucesión creciente $a_n \in A$ tal que $a_n \to s$. Si $s \notin A$, se puede elegir a_n estrictamente decreciente.*

Demostración. Si $s \in A$, basta tomar $a_n = s$ la sucesión constante (que, por definición es creciente). Si no, tomamos algún elemento $\overline{a_1} \in A$, y ponemos $a_1 := \overline{a_1}$.

Ahora consideremos $\overline{a_2} = \frac{a_1+s}{2}$. Vale que $\overline{a_2} < s$. Entonces existe $a_2 \in A$ tal que $\overline{a_2} < a_2 < s$. Si no existiera, $\overline{a_2}$ sería cota superior de A, lo cual es imposible pues $\overline{a_2} < s$. Así vamos definiendo $\overline{a_n}$ como el promedio $\frac{a_{n-1}+s}{2}$, y tomamos como a_n cualquier elemento de A que está entre $\overline{a_n}$ y s. La sucesión a_n verifica ser creciente (pues $a_n > \overline{a_n} > a_n - 1$) y vale que

$$0 < a_1 < a_2 < a_3 < \ldots < a_n$$

$$0 < s - a_n < \tfrac{s-a_{n-1}}{2} < \ldots < \tfrac{s-a_1}{2^{n-1}}$$

O sea que $\lim\limits_{n\to\infty} a_n = s$. □

Proposición 4. *Toda sucesión $\{a_n\} \subseteq \mathbb{R}$ tiene una subsucesión $\{a_{n_k}\}$ monótona.*

Demostración. Se dice que un término a_i es **dominante** si verifica que $\forall k \geq i$ vale que $a_k \leq a_i$. Ahora se deben considerar dos casos: que haya finitos o infinitos términos dominantes:

- **Infinitos términos dominantes:** Sean $n_1 < n_2 < n_3, \ldots$ los índices tales que a_{n_i} es un término dominante. La subsucesión $\{a_{n_i}\}$, por tener todos términos dominantes, verifica que:

  $$a_{n_1} \geq a_{n_2} \geq a_{n_3} \geq \ldots$$

  Con lo que hemos construido una subsucesión decreciente y, por ende, monótona.

- **Finitos términos dominantes:** Sea a_u el último término dominante. Consideremos $a_{n_1} := a_{u+1}$ o sea, el primer término después del último término dominante (si no hay ningún término dominante tomamos $n_1 = 1$). Como a_{n_1} no es dominante, sabemos que existe n_2 tal que $a_{n_2} > a_{n_1}$. A su vez, como a_{n_2} tampoco es dominante (pues $n_2 > u$ y u era el último término dominante), tiene que existir un n_3 tal que $a_{n_3} > a_{n_2} > a_{n_1}$. De esa forma, siempre podemos hallar n_k tal que

  $$a_{n_k} > a_{n_{k-1}} > \ldots > a_{n_2} > a_{n_1}$$

  Con esto, hemos construido una subsucesión creciente y, por ende, monótona.

□

Proposición 5 (Bolzano-Weierstrass)**.** *Toda sucesión $\{a_n\} \subseteq \mathbb{R}$ acotada tiene una subsucesión $\{a_{n_k}\}$ convergente.*

Demostración. Sea $\{a_{n_k}\}$ una subsucesión monótona de $\{a_n\}$ (que se puede extraer por la proposición anterior). Como $\{a_{n_k}\}$ es una subsucesión monótona y acotada (por ser $\{a_n\}$ acotada), contiene una subsucesión $\{a_{n_{k_j}}\} = \{a_{n_{k'}}\}$ convergente. □

Proposición 6 (Cauchy)**.** *Sea a_k una sucesión de números reales y sea $S_n = \sum\limits_{k=0}^{n} a_k$ la sucesión de sumas parciales. Entonces*

1. *$a_k \to 0$ no implica que la serie converge.*

2. *Si la serie converge, entonces $a_k \to 0$.*

3. *Una serie de términos positivos $a_k \geq 10$ es convergente si y sólo si sus sumas parciales están acotadas, pues en este caso, S_n es una sucesión creciente de números reales.*

3 Conjuntos

Proposición 7. *Un conjunto $C \subset \mathbb{R}^n$ es cerrado si y sólo si su complemento C^c es abierto.*

Demostración.
$\Rightarrow$) Veamos que C^c es abierto, o sea, que $\forall p \in C^c,\ p$ es interior. Sea $p \in C^c$ no interior. Entonces, $\forall k \in \mathbb{N}$ se tiene $B_{\frac{1}{k}}(P) \nsubseteq C^c$, o sea, existe $P_k \in C^{c^c} = C$ tal que $P_k \in B_{\frac{1}{k}}(P)$. Observemos que $P_k \to P$ (pues $||P_k - P|| < \frac{1}{k}$). Hemos llegado a una contradicción pues $P_k \in C$, $P \notin C$ y, por hipótesis, C es cerrado.

$\Leftarrow$) Consideremos $\{P_k\} \subseteq C$ una sucesión de puntos de C con límite $P \in \mathbb{R}^n$. Queremos ver que $P \in C$ o, equivalentemente, $P \notin C^c$. Si se diera lo segundo, existiría $r > 0$ tal que $B_r(P) \subset C^c$. Consideremos $k \in \mathbb{N}$ tal que $||P_k - P|| < r$. Entonces, tendríamos que $P_k \in B_r(P) \subset C^c$, lo que es imposible pues $P_k \in C$. Luego, $P \in C$, lo que demuestra que C es cerrado. $\square$

Proposición 8. *La clausura $\overline{C}$ es un conjunto cerrado.*

Demostración. Sea $P \in \overline{C}^c$ un punto cualquiera. Entonces afirmo que existe $r > 0$ tal que $B_r(P) \bigcap C = \emptyset$ (de lo contrario, uno podría fabricar una sucesión X_k de puntos de C tal que $X_k \to P$, lo que diría que $P \in \overline{C}$, contradiciendo que $P \in \overline{C}^c$).

Esto implica que P es un punto interior de $\overline{C}^c$ y, como no impusimos ninguna restricción sobre P, vale que $\forall P \in \overline{C}^c$, P es interior. Esto implica que $\overline{C}^c$ es un conjunto abierto, demostrando que $\overline{C}$ es un conjunto cerrado. $\square$

4 Funciones

Proposición 9. *Sea $F : A \subset \mathbb{R}^n \to \mathbb{R}^m$. Sean $P \in \overline{A}$ y $L \in \mathbb{R}^m$. Las siguientes afirmaciones son equivalentes:*

- 1) $\lim\limits_{X \to P} F(X) = L$.
- 2) *Para* ***toda*** *sucesión de puntos $P_k \to P$, se tiene que $\lim\limits_{n \to \infty} F(P_k) = L$.*

Demostración.
(1) $\Rightarrow$ (2) Sea P_k una sucesión de puntos de A que tiende a P. Entonces por la definición del límite de función, dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$0 < ||X - P|| < \delta \text{ y } X \in A \Rightarrow ||F(X) - L|| < \varepsilon.$$

Si tomamos $k_0 \in \mathbb{N}$ tal que $k \geq k_0$ implique $||P_k - P|| < \delta$, lo anterior nos garantiza que $||F(P_k) - L|| < \varepsilon$. O sea $F(P_k) \to L$ de acuerdo a la definición de límite de sucesiones, aplicado a la sucesión $Q_k = F(P_k)$.

(1) $\Leftarrow$ (2) Negar que $\lim\limits_{X \to P} F(X) = L$ equivale a decir que existe $\varepsilon > 0$ tal que para todo $\delta > 0$ existe $X \in A$ con $0 < ||X - P|| < \delta$ y $||F(X) - L|| \geq \varepsilon$. Tomemos $\delta = \frac{1}{k}$, con $k \in \mathbb{N}$, y sea P_k el punto correspondiente. Vemos que $P_k \to P$ pero, por otro lado, $||F(P_k) - L|| \geq \varepsilon$, con lo cual $F(P_k) \nrightarrow L$, contradiciendo la hipótesis. ☐

Proposición 10 (Bolzano)**.** *Si $f : [a, b] \to \mathbb{R}$ es continua y $f(a).f(b) < 0$, entonces existe $c \in (a, b)$ tal que $f(c) = 0$.*

Demostración. Supongamos $f(a) > 0$ y $f(b) < 0$. Sea $A = \{x \in [a, b] : f(x) > 0\} \neq \emptyset$ (pues $a \in A$). A tiene supremo por estar acotado superiormente (pues $A \subset [a, b]$) y sea $s = \sup A$. Afirmo entonces que $f(s) = 0$. Si no fuera así, debería ser $f(s) > 0$ o $f(s) < 0$.

- $f(s) > 0$. Debería existir un entorno $(s - \varepsilon, s + \varepsilon)$ al rededor de s tal que para todo x en el entorno $f(x) > 0$. Sea un punto $x_0 \in (s, s + \varepsilon)$. Este x_0 verifica que $x_0 \in A$ (porque $f(x_0) > 0$ y $x_0 < b$) y $x_0 > s$ lo cual es absurdo pues s es el supremo de A.
- $f(s) < 0$. Se demuestra en forma análoga al caso anterior que es absurdo.

Ergo, la única opción que queda es que $f(s) = 0$, que es lo que queríamos demostrar. ☐

Demostración alternativa. Supongamos $f(a) > 0$ y $f(b) < 0$. Sea $A = \{x \in [a, b] : f(x) > 0\} \neq \emptyset$ (pues $a \in A$). A tiene supremo por estar acotado superiormente (pues $A \subset [a, b]$) y sea $s = \sup A$. Entonces existe una sucesión creciente $\{a_n\}$ de puntos de A que tiende a $s \in [a, b]$.

Como $f(a_n) > 0$ (pues $a_n \in A$) entonces $f(s) \geq 0$. Afirmo entonces que $f(s) = 0$. Si no fuera así, debería ser $f(s) > 0$ y eso es imposible porque debería existir un entorno $(s - \varepsilon, s + \varepsilon)$ al rededor de s tal que $f(x) > 0$. Pero, como $s < b$, podemos tomar $x_0 \in [a, b]$ tal que $x_0 \in (s, s + \varepsilon)$. Este $x_0 \in A$ y es mayor que el supremo, lo que es absurdo. O sea $f(s) = 0$. ☐

Proposición 11 (Weiestrass)**.** *Sean $A \subset \mathbb{R}^n$ compacto y $f : A \to \mathbb{R}$ continua en A. Entonces vale que:*

- 1) *Existen $m, M \in \mathbb{R}$ tales que $m \leq f(X) \leq M$ para todo $X \in A$.*
- 2) *Existen $P_m, P_M \in A$ tales que $f(P_m) = min\{f(X) : X \in A\}$ y $f(P_M) = max\{f(X) : X \in A\}$. O sea, f alcanza su máximo y mínimo en A.*

Demostración.
1) Supongamos que f no es acotada, o sea, que no existe $M \in \mathbb{R}$ tal que $f(X) \leq M$. Entonces debe existir una sucesión de puntos $\{X_k\} \in A$ tal que $f(X_k) \geq k$ para todo $k \in \mathbb{N}$. Como A es cerrado y acotado, $\{X_k\}$ necesariamente tiene una subsucesión $\{X_{k_j}\}$ que converge a $P \in A$. (Esto es porque toda sucesión en un conjunto acotado tiene límite en la clausura del conjunto. Como, en este caso A es cerrado, su clausura

coincide consigo mismo).
Sin embargo, como $f(X_k) \geq k$ y f es continua, es imposible que exista $\{X_{k_j}\}$ convergente a P.

2) Por lo anterior, vale que $Im(f) \subseteq [m, M]$, o sea, es un conjunto acotado de $\mathbb{R}$. Sea $s = sup(Im(f))$. Queremos ver que en realidad s es un máximo, o sea, que existe $P_M \in A$ tal que $f(P_M) = s$. Consideremos la sucesión $\{X_k\} \in A$ que, aplicándole la función f, tienda al supremo. Tenemos entonces una sucesión $\{X_k\} \in A$ tal que $\lim_{k\to\infty} f(X_k) = s$.
Ahora, de $\{X_k\}$ extraemos una subsucesión $\{X_{k_j}\}$ convergente y llamamos a su límite P_M (que, por ser $\{X_k\}$ una sucesión de A y este ser cerrado, $P_M \in A$). Como f es continua $f(P_M) = \lim_{k\to\infty} f(X_k) = s$. □

Proposición 12 (Heine-Cantor)**.** *Continua en un compacto implica uniformemente continua. O sea, sea $F : A \subseteq R^n \to R^m$ continua con A compacto. Entonces, vale que F es uniformemente continua.*

Demostración. Supongamos que f es continua pero no *uniformemente* continua. Entonces, debe existir $\varepsilon_0 > 0$ tal que para todo $\delta > 0$, existan $x, y \in A$ que verifiquen que $||F(x) - F(y)|| \geq \varepsilon_0$.

Consideremos dos sucesiones $\{X_k\}, \{Y_k\} \in A$ tales que $||X_k - Y_k|| < \frac{1}{k}$. Por lo anteriormente dicho, debe valer que $||F(X_k) - F(Y_k)|| \geq \varepsilon_0$. Como A es compacto, deben existir dos subsucesiones $\{X_{k_j}\}, \{Y_{k_j}\}$ convergentes a X_0 e Y_0 respectivamente. Consideremos entonces

$$||Y_{k_j} - X_0|| = ||Y_{k_j} - X_{k_j} + X_{k_j} - X_0|| \leq ||Y_{k_j} - X_{k_j}|| + ||X_{k_j} - X_0|| < \frac{1}{k_j} + ||X_{k_j} - X_0||$$

O sea que ambas sucesiones convergen al mismo valor $X_0 = Y_0$. Sin embargo, esto no es posible pues contradice que $||F(X_k) - F(Y_k)|| \geq \varepsilon_0$. □

Proposición 13. *Sea $G : [a, b) \to \mathbb{R}$ una función creciente (b puede ser $+\infty$). Entonces existe el límite $\lim_{x\to b^+} G(x)$. Además, este límite es finito si y sólo si G es acotada superiormente.*

Demostración. Sea $l = \sup\{G(x) : x \in [a, b)\}$. Por definición de supremo, dado $\varepsilon > 0$, existe $x_0 \in [a, b)$ tal que $l - G(x_0) < \varepsilon$. Si $|b - x| = b - x < \delta = b - x_0$, entonces $x \leq x_0$ y G es creciente $G(x) \leq G(x_0)$. Entonces

$$l - G(x) \leq l - G(x_0) < \varepsilon$$

lo que prueba que el límite existe y coincide con el supremo. En particular el límite es finito si y sólo si G es acotada. Entonces, $l = \lim_{x\to b^+} G(x)$ □

5 Diferenciación

Proposición 14 (Diferenciable ⇒ Continua)**.** *Sea $f : A \subset \mathbb{R}^n \to R$, $P \in A^o$. Si f es **diferenciable** en P, entonces f es **continua** en P.*

Demostración. Consideremos

$$\begin{aligned} |f(X) - f(P)| &\leq |f(X) - f(P) - Df_p(X-P) + Df_p(X-P)| \\ &\leq |f(X) - f(P) - Df_p(X-P)| + |Df_p(X-P)| \\ &\leq |f(X) - f(P) - Df_p(X-P)| + ||\nabla f_p||\ ||(X-P)|| \end{aligned}$$

El último término sale por la desigualdad de Cauchy-Schwarz

$$|Df_p(X)| = |\langle \nabla f_p, X \rangle| \leq ||\nabla f_p||\ ||X||$$

Dividiendo y multiplicando por $||X - P||$ el primer sumando obtenemos:

$$|f(X) - f(P)| \leq \frac{|f(X) - f(P) - Df_p(X-P)|}{||X-P||}\ ||X-P|| + ||\nabla f_p||\ ||X-P|| \leq 0$$

El primer sumando de la derecha tiende a cero cuando $X \to P$ por ser f diferenciable en P. El segundo trivialmente también tiende a cero, con lo cual obtenemos que $\lim\limits_{X \to P} f(X) = f(P)$ o, equivalentemente, f es continua en P. □

Proposición 15 (Unicidad del Diferencial)**.** *Sea $f : A \subset \mathbb{R}^n \to R$, $P \in A^o$. Si existe una transformación lineal $T_p : \mathbb{R}^n \to \mathbb{R}$ tal que*

$$\lim_{X \to P} \frac{|f(X) - f(P) - T_p(X-P)|}{||X-P||} = 0 \tag{1}$$

Entonces
1) *Existen todas las derivadas direccionales de f en P y vale que*

$$\frac{\partial f}{\partial V}(P) = T_p(V) \text{ para todo } V \in \mathbb{R}^n,\ ||V|| = 1$$

2) *En particular, existen todas las derivadas parciales de f, se tiene $f_{x_i}(P) = T_P(E_i)$ y la transformación T_P es única.*
3) *Se tiene $T_P(X) = Df_P(X) = \langle \nabla f_P, X \rangle$ para todo $X \in \mathbb{R}^n$, y f es diferenciable en P. En particular, si $||V|| = 1$,*

$$\frac{\partial f}{\partial V}(P) = \langle \nabla f_P, V \rangle = Df_P(V)$$

Demostración.
1) Sea $V \in \mathbb{R}^n$ tal que $||V|| = 1$. Como asumimos que existe el límite (1), también existe compuesto con cualquier curva que tenga límite P. En particular, si ponemos $X = P + tV$ (con t suficientemente chico para que $X \in A$), vale que $X - P = tV$. Tomando módulo, $||X - P|| = |t|$, lo que implica que

$$0 = \lim_{t \to 0} \frac{|f(P+tV) - f(P) - T_P(P+tV-P)|}{|t|} = \lim_{t \to 0} \left| \frac{f(P+tV) - f(P)}{t} - T_P(V) \right|$$

Esto prueba que $f_V(P) = \lim\limits_{t \to 0} \dfrac{f(P+tV) - f(P)}{t} = T_P(V)$. Como V puede ser igual a E_i, también prueba que existen todas las derivadas parciales.

Además, como $T_P(E_i) = f_{x_i}(P)$ vale para toda E_i y una transformación lineal queda determinada por su valor en una base de $\mathbb{R}^n$, se deduce que $T_P(X) = \langle \nabla f_P, X \rangle$.

Supongamos que existe otra transformación T'_P que verifique la condición del límite. Entonces, también debe cumplir que $T'_P(X) = \langle \nabla f_P, X\rangle = T_P(X)$.

$\square$

Proposición 16. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ y $P \in A$. El gradiente (∇f_P) es la dirección de mayor crecimiento de f en P.*

Demostración. Sea V de norma unitaria.

$$f_V(P) = \langle \nabla f_P, V\rangle \leq |\langle \nabla f_P, V\rangle| \leq ||\nabla f_P||\ ||V|| = ||\nabla f_P||$$

Esto demuestra que la derivada direccional a lo sumo vale $||\nabla f_P||$. Por otro lado, si consideramos la dirección del gradiente normalizada poniendo $V_P = \frac{\nabla f_P}{||\nabla f_P||}$, se tiene:

$$f_{V_P} = \langle \nabla f_P, V_P\rangle = \langle \nabla f_P, \tfrac{\nabla f_P}{||\nabla f_P||}\rangle = ||\nabla f_P||$$

$\square$

Teorema 17 (Fermat)**.** *Si $f : [a, b] \to \mathbb{R}$ es derivable en $c \in (a, b)$ y c es un extremo local de f, entonces $f'(c) = 0$.*

Demostración. Supongamos que c es máximo local de f. Por definición, esto significa que existe $\varepsilon > 0$ tal que si $x \in (c - \varepsilon, c + \varepsilon)$ entonces $f(x) \leq f(c)$. Calculemos $f'(c)$ por definición, usando límites laterales:

$$f'(c) = \lim_{x\to 0^-} \frac{f(c + h) - f(c)}{h} \geq 0$$

Esto se debe a que $h < 0$ (pues tiende a 0 por izquierda) y $f(c + h) - f(c) \leq 0$ (pues c es máximo local por hipótesis).

Análogamente considerando el límite por derecha tenemos que:

$$f'(c) = \lim_{x\to 0^+} \frac{f(c + h) - f(c)}{h} \leq 0$$

Juntando ambas ecuaciones, debe valer que $0 \leq f'(c) \leq 0$, lo que implica $f'(c) = 0$ $\square$

Teorema 18 (Rolle)**.** *Si $f : [a, b] \to \mathbb{R}$ es continua en $[a, b]$ y derivable en (a, b) y además $f(a) = f(b)$, entonces existe $c \in (a, b)$ tal que $f'(c) = 0$.*

Demostración. Si f es continua en $[a, b]$ y este último conjunto es compacto, f alcanza su máximo y mínimo en $[a, b]$. Separemos en dos casos:

- $max = min$. Si este es el caso, la función es la función constante y tomando cualquier $c \in (a, b)$ vale que $f'(c) = 0$.

- $max \neq min$. Como $f(a) = f(b)$, puede ser que $f(a) = f(b) = max$ o $f(a) = f(b) = min$ (o ninguna de las dos). Pero necesariamente va a existir $c \in (a, b)$ tal que c es máximo o mínimo; o sea c es extremo local. Por el teorema de *Fermat*, $f'(c) = 0$

$\square$

Teorema 19 (Lagrange)**.** *Si $f : [a, b] \to \mathbb{R}$ es continua en $[a, b]$ y derivable en (a, b) entonces existe $c \in (a, b)$ tal que*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Demostración. Sea la función $L(x)$ la recta que pasa por los puntos $(a, f(a))$ y $(b, f(b))$. O sea,

$$L(x) = (x-a)\frac{f(b)-f(a)}{b-a} + f(a)$$

Consideremos la función auxiliar $g(x) = f(x) - L(x)$. Dado que $L(a) = f(a)$ y $L(b) = f(b)$, vale que $g(a) = g(b) = 0$. Además, por álgebra de funciones continuas y derivables, vale que g es continua en $[a,b]$ y derivable en (a,b). Entonces g cumple las hipótesis del teorema de Rolle, lo que nos permite afirmar que existe $c \in (a,b)$ tal que $g'(c) = 0$.

Sea m la pendiente de la recta L. Entonces, tenemos que

$$0 = g'(c) = f'(c) - L'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}.$$

Reagrupando apropiadamente, tenemos que $f'(c) = \dfrac{f(b)-f(a)}{b-a}$ □

Teorema 20 (Cauchy)**.** *Si $f, g : [a,b] \to \mathbb{R}$ son continuas en $[a,b]$ y derivables en (a,b), entonces existe $c \in (a,b)$ tal que*

$$f'(c)(g(b) - g(a)) = g'(c)(f(b) - f(a))$$

Demostración. Consideremos la función auxiliar

$$h(x) = (f(x) - f(a))(g(b) - g(a)) \;-\; (g(x) - g(a))(f(b) - f(a))$$

Dado que $h(a) = h(b) = 0$ y, por construcción, es continua en $[a,b]$ y derivable en (a,b). Consecuentemente h cumple las hipótesis del teorema de Rolle. O sea existe $c \in (a,b)$ tal que $h'(c) = 0$. Pero, por otra parte, vale que

$$0 = h'(c) = f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a))$$

Despejando apropiadamente se obtiene la tesis del teorema. □

Observación 21. *Sea $F(u,v) = (f_1(u,v), f_2(u,v), f_3(u,v))$. Entonces, la fórmula de la derivada parcial respecto a la primer coordenada de la composición queda escrita de la siguiente manera:*

$$\frac{\partial(g \circ F)}{\partial u} = \frac{\partial g}{\partial f_1}\frac{\partial f_1}{\partial u} + \frac{\partial g}{\partial f_2}\frac{\partial f_2}{\partial u} + \frac{\partial g}{\partial f_3}\frac{\partial f_3}{\partial u}$$

Teorema 22 (Lagrange en $\mathbb{R}^n$)**.** *Sea $f : A \subseteq \mathbb{R}^n \to R$ diferenciable con A* ***convexo*** *(o sea, para todo par de puntos del conjunto el segmento que los une está totalmente contenido en el conjunto). Entonces, para todo $Q, R \in A$, existe P_0 en el segmento que une Q con R tal que*

$$f(Q) - f(R) = \langle D_f(P_0), Q - R \rangle$$

Demostración. Consideremos la parametrización del segmento que une Q con R, $\alpha(t) = R + t(Q - R)$. Entonces, $h(t) = f \circ \alpha(t)$ está definida y es continua en $[0,1]$ y es derivable en $(0,1)$ (por ser composición de continuas/diferenciables). Entonces, por el teorema de *Lagrange* en una variable, existe $c \in (0,1)$ tal que $h(1) - h(0) = h'(c)$. Pero, por regla de la cadena, $h'(c) = Df_{\alpha(c)}\alpha'(c)^t = Df_{g'(c)}(Q-R)^t = \langle D_f(\alpha(c)), Q - R \rangle$. Llamando P_0 a $\alpha(c)$ obtenemos la tesis. □

Proposición 23. *Si $G : B_r(P) \subset \mathbb{R}^n \to \mathbb{R}^n$ es diferenciable y $||D_G Q||_\infty \leq M$ para todo $Q \in B_r(P)$, entonces se tiene que $||G(X) - G(Y)|| \leq M||X - Y||$ para todo $X, Y \in B_r(P)$*

Demostración. Si $G(X) = G(Y)$, entonces $0 \leq M||X - Y||$ es trivial. Supongamos $G(X) \neq G(Y)$. Consideremos el segmento que une X con Y, o sea el segmento $g(t) = Y + t(X - Y)$. Pongamos la función auxiliar:

$$h(t) = \langle G \circ g(t), G(X) - G(Y) \rangle$$

$$h'(t) = \langle DG_{g(t)}g'(t), G(X) - G(Y)\rangle$$

$h'(t)$ verifica las hipótesis del teorema de Lagrange, lo que nos permite afirmar que existe $c \in (0,1)$ tal que

$$h(1) - h(0) = h'(c)(1 - 0)$$

Reemplazando por la definición de h, esto nos lleva a la expresión:

$$\langle G(X) - G(Y), G(X) - G(Y)\rangle = \langle DG_{Y+c(X-Y)(X-Y)}, G(X) - G(Y)\rangle$$
$$||G(X) - G(Y)||^2 \leq ||DG_{Y+c(X-Y)}(X - Y)||_\infty ||G(X) - G(Y)|| \leq M||X - Y||\ ||G(X) - G(Y)||$$

Como supusimos $G(X) \neq G(Y)$, entonces $||G(X) - G(Y)|| \neq 0$, con lo cual podemos dividir ambos miembros por ese término, demostrando la afirmación buscada. $\square$

6 Implícita e Inversa

Observación 24. *Si F es diferenciable e inversible, y su inversa también es diferenciable, entonces por regla de la cadena aplicada a $F^{-1}(F(X))$ vale que*

$$DF^{-1}_{F(X)} DF_X = Id$$

Donde Id es la transformación identidad de $\mathbb{R}^n$. De aquí se deduce que la diferencial de la inversa es la inversa de la diferencial de F, o sea:

$$DF^{-1}_{F(X)} = (DF_X)^{-1}$$

Proposición 25. *Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal tal que $||I - T||_\infty < 1$. Entonces T es inversible.*

Demostración. Para ver que es inversible alcanza con ver que $ker(T) = \{\mathbb{O}\}$. Supongamos que no lo es. O sea, que existe $V \neq \mathbb{O}$ tal que $TV = \mathbb{O}$. Pero entonces,

$$||V|| = ||V - \mathbb{O}|| = ||V - TV|| = ||(1 - T)V|| \leq ||(1 - T)||_\infty \; ||V|| < 1||V|| = ||V||$$

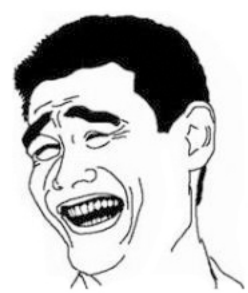
Concluimos que $||V|| < ||V||$ que es claramente un absurdo. O sea, necesariamente $ket(T) = \{\mathbb{O}\}$ lo que implica que T es inversible. $\square$

Teorema 26 (Función Inversa)**.** *Sean $F : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ con A abierto y $P \in A$. Si F es $\mathbb{C}^1$ en A y $DF_P : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal inversible, entonces existen entornos abiertos V, W de P y $F(P)$ respectivamente tales que $F : V \to W$ es biyectiva, la inversa F^{-1} es diferenciable y la además vale que, para todo $Q \in V$*

$$DF^{-1}_F(Q) = (DF_Q)^{-1}$$

Con lo cual, en particular, la diferencial de F es inversible en V.

Demostración.



$\square$

Teorema 27 (Función Implícita en $\mathbb{R}^2$)**.** *Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ una función C^1 y $S = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$ una curva de nivel c de f. Supongamos que $\dfrac{\partial f}{\partial y}(P) \neq 0$ para algún $P \in S$. Entonces:*

- 1. *Existen un intervalo abierto $I \subset \mathbb{R}$, una función derivable $\varphi : I \to \mathbb{R}$ y una bola B al rededor de P tales que $S \bigcap B = Gr(\varphi)$*
- 2. $\dfrac{\partial f}{\partial y}(x, y) \neq 0$ *para todo $(x, y) \in S \bigcap B$ y $\nabla f_{(x,y)}$ es perpendicular a S en $S \bigcap B$.*
- 3. *Para todo $x \in I$ se tiene que*

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}(x, \varphi(x))$$

Demostración.

$\square$

Observación 28. *Sea $S = \{(x, y) : f(x, y) = k\}$ y sea $\alpha(t)$ una parametrización de la curva S. Entonces, $(f \circ \alpha)(t) = k$ para todo t en el dominio de α. Si asumimos que todas las funciones son derivables, por la regla de la cadena, vale que*

$$(f \circ \alpha)'(t) = \nabla f(\alpha(t))(\alpha'(t)) = 0$$

O sea que $\nabla f_{(x,y)}$ es la dirección de la recta normal a la curva de nivel en el punto $(x, y) \in S$

Teorema 29. *Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ una función C^1 y $S = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ una curva de nivel de f. Supongamos que $\frac{\partial f}{\partial y}(P) \neq 0$ para algún $P \in S$. Entonces*

- 1) *Existen un intervalo abierto $I \subset \mathbb{R}$, una función derivable $\varphi(x), \varphi : I \to \mathbb{R}$ y una bola B al rededor de P tales que $S \bigcap B = Gr(\varphi)$*
- 2) $\frac{\partial f}{\partial y}(x, y) \neq 0$ *para todo $(x, y) \in S \bigcap B$. Además $\nabla f_{(x,y)}$ es perpendicular a S en $S \bigcap B$.*
- 3) *Para todo $x \in I$ se tiene que*

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}(x, \varphi(x))$$

7 Taylor, extremos y Lagrange

Proposición 30. *Sea I un intervalo abierto en $\mathbb{R}$ y sea $f \in C^n(I)$. El* ***polinomio de Taylor*** *de grado n de f en el punto $a \in I^o$ es el único polinomio $P(x)$ de grado n que, para todo $k \in \{0, 1, \ldots, n\}$, verifica*

$$P^{(k)}(a) = f^{(k)}(a)$$

La expresión de P es la siguiente:

$$P(x) = f(a) = f'(a)(x-a) + \tfrac{1}{2} f''(a)(x-a)^2 + \ldots + \tfrac{1}{n!} f^{(n)}(a)(x-a)^n$$

Además, para todo $x \in I$ vale que:

$$f(x) = P(x) + R(x)$$

donde $R(x) = f(x) - P(x)$ es el resto, que verifica que $\lim\limits_{x \to a} \dfrac{R(x)}{(x-a)^n} = 0$.

Teorema 31 (Taylor con resto de Lagrange)**.** *Sea I un intervalo abierto en $\mathbb{R}$ y sea $f : I \to \mathbb{R}$ una función $n+1$ veces derivable. Entonces, dados $x, a \in I$, existe c estrictamente entre x y a tal que*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

Demostración. Fijados x, $a \in I$, consideremos la siguiente función auxiliar dada por $g : I \to \mathbb{R}$:

$$g(t) = f(x) - f(t) - f'(t)(x-t) - \tfrac{1}{2} f''(t)(x-t)^2 - \ldots - \tfrac{1}{n!} f^{(n)}(t)(x-t)^n - \tfrac{K}{(n+1)!}(x-t)^{n+1}$$

Siendo K una constante apropiada tal que $g(a) = 0$. Intentemos ver cuánto vale dicha constante. Observemos que $g(t) = f(x) - P_f(t)$ (siendo $P_f(t)$ el polinomio de Taylor de f centrado en t). En consecuencia, g es una función derivable de la variable t y continua en el intervalo cerrado entre x y a. Además, se tiene que $g(x) = g(a) = 0$, con lo cual, por el teorema de Rolle, existe c entre x y a tal que $g'(c) = 0$. Pero, se tiene que

$$0 = g'(c) = \frac{K - f^{(n+1)}(c)}{n!}(x-t)^n$$

$$\Rightarrow K = f^{(n+1)}(c)$$

Lo único restante es evaluar g en el punto a y despejar el valor de $f(x)$. □

Teorema 32 (Clairaut-Schwarz)**.** *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A abierto. Si $f \in C^2(A)$ entonces las derivadas cruzadas coinciden. Es decir, para todos i, $j \in \{1, 2, \ldots, n\}$ se tiene que*

$$\frac{\partial^2 f}{\partial x_i \, \partial x_j}(P) = \frac{\partial^2 f}{\partial x_j \, \partial x_i}(P)$$

Demostración. Para simplificar la demostración, consideremos una función $f : \mathbb{R}^2 \to \mathbb{R}$ y un punto $P = (a, b) \in A$. A su vez, consideremos las funciones auxiliares

$$g(t) = f(a+t, b+t) - f(a+t, b) - f(a, b+t) + f(a, b)$$

$$\varphi(x) = f(x, b+t) - f(x, b)$$

Dado que la función φ es dos veces derivable, podemos plantear el polinomio de Taylor de $\varphi(a+t)$ de la siguiente forma:

$$\varphi(a+t) = \varphi(a) + \varphi'(a)t + \varphi''(c)\tfrac{t^2}{2}$$

donde $\varphi''(c)\frac{t^2}{2}$ es el resto de Lagrange (lo que permite poner la igualdad). Restando a ambos miembros $\varphi(a)$, obtenemos que

$$\varphi'(a)t + \varphi''(c)\tfrac{t^2}{2} = \varphi(a+t) - \varphi(a) = g'(t)$$

Reemplazando por la definición de φ, obtenemos

$$f'(a, b+t)t + f''(c, b+t)\tfrac{t^2}{2} =$$

$$= \left(\tfrac{\partial f}{\partial x}(a, b+t) - \tfrac{\partial f}{\partial x}(a, b)\right) t + \left(\tfrac{\partial^2 f}{\partial x\, \partial x}(c, b+t) - \tfrac{\partial^2 f}{\partial x\, \partial x}(c, b)\right) \tfrac{t^2}{2}$$

Si dividimos por t^2 y hacemos tender t a 0, obtenemos

$$= \frac{\left(\tfrac{\partial f}{\partial x}(a, b+t) - \tfrac{\partial f}{\partial x}(a, b)\right) \cancel{t}}{t^{\cancel{2}}} + \frac{\left(\tfrac{\partial^2 f}{\partial x\, \partial x}(c, b+t) - \tfrac{\partial^2 f}{\partial x\, \partial x}(c, b)\right) \tfrac{\cancel{t^2}}{2}}{\cancel{t^2}}$$

Como $\frac{\partial^2 f}{\partial x\, \partial x}$ es continua, el último término tiende a cero. O sea

$$\frac{g(t)}{t^2} = \lim_{t \to 0} \frac{\left(\tfrac{\partial f}{\partial x}(a, b+t) - \tfrac{\partial f}{\partial x}(a, b)\right)}{t}$$

Por definición, esto es la derivada con respecto a y de la función derivada con respecto a x. O sea:

$$\lim_{t \to 0} \frac{g(t)}{t^2} = \frac{\partial^2 f}{\partial y\, \partial x}(a, b)$$

Ahora podemos repetir el argumento considerando $\psi(y) = f(a+t, y) - f(a, y)$. Se tiene entonces que

$$g(t) = \psi(b+t) - \psi(b) = \psi'(b)t + \psi''(c)\tfrac{t}{2}$$

para algún otro c entre b y $b+t$. Escribiendo las derivadas y razonando como antes, vale que

$$\lim_{t \to 0} \frac{g(t)}{t^2} = \frac{\partial^2 f}{\partial x\, \partial y}(a, b)$$

$\square$

Teorema 33 (Taylor en $\mathbb{R}^2$)**.** *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A abierto convexo. Supongamos que f es C^3 en A. Entonces, dado $P = (p_0, p_1, \ldots, p_n) \in A$, para todo $X = (x_1, x_2, \ldots, x_n) \in A$ se tiene que*

$$f(X) = f(P) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i - p_i) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j\, \partial x_i}(P)(x_j - p_j)(x_i - p_i) + R_p(X - P)$$

Donde la expresión del resto de Lagrange es:

$$R_P(X - P) = \frac{1}{6} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_k\, \partial x_j\, \partial x_i}(C)(x_k - p_k)(x_j - p_j)(x_i - p_i)$$

con C algún punto en el segmento entre X y P y verifica que:

$$\lim_{X \to P} \frac{R_P(X - P)}{||X - P||^2} = 0$$

Demostración. Consideremos X, $P \in A$ y consideremos la función auxiliar

$$g(t) = f(P + t(X - P))$$

que es C^3 en un entorno del intervalo $[0, 1]$. Derivando, obtenemos

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(P + t(X - P))(x_i - p_i)$$

Derivando nuevamente, obtenemos

$$\begin{aligned} g''(t) &= \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}(P+t(X-P))\right)'(x_i-p_i)\\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_j}(P+t(X-P))(x_j-p_j)\right)(x_i-p_i)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_j\,\partial x_i}(P+t(X-P))(x_j-p_j)(x_i-p_i) \end{aligned}$$

Derivamos una vez más y obtenemos

$$\begin{aligned} g'''(t) &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial^2 f}{\partial x_j\,\partial x_i}(P+t(X-P))\right)'(x_j-p_j)(x_i-p_i)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^3 f}{\partial x_k\,\partial x_j\,\partial x_i}(P+t(X-P))(x_k-p_k)(x_j-p_j)(x_i-p_i) \end{aligned}$$

Ahora utilizamos la fórmula de Taylor de orden dos, en una variable g en el intervalo $[0,1]$, que nos dice que

$$g(1)=g(0)+g'(0)+\frac{1}{2}g''(0)+\frac{1}{6}g'''(c)$$

donde $c\in(0,1)$. Observemos que $C=P+c(X-P)$ es, en efecto, un punto en el segmento que une X con P. Reemplazando en esta fórmula la definición de g y sus derivadas se tiene la fórmula del enunciado del teorema.
Sin embargo es necesario ver que el límite da cero. Para eso, debemos observar que, dado que las derivadas terceras son funciones continuas (pues la función es C^3), tomando algún compacto al rededor de P, son funciones acotadas. Por esto, si X está cerca de P, se tiene que

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left|\frac{\partial^3 f}{\partial x_k\,\partial x_j\,\partial x_i}(C)\right|\le M$$

A su vez, como $|x_k-p_k||x_j-p_j||x_i-p_i|\le ||X-P||^3$, se deduce que

$$|R_P(X-P)|\le\frac{1}{6}M||X-P||^3$$

si X está suficientemente cerca de P. Dividiendo por $||X-P||^2$ y tomando límite $X\to P$ se tiene la desigualdad buscada. ☐

También es posible escribir esta misma demostración con notación vectorial

Demostración con notación vectorial. Consideremos la misma función auxiliar $g(t)=f(P+t(X-P))$ y sus respectivas derivadas:

$$g'(t)=\langle\nabla f_{P+t(X-P)},X-P\rangle$$

$$g''(t)=\langle(\nabla f_{P+t(X-P)})',X-P\rangle=\langle Hf_{P+t(X-P)}\cdot(X-P),X-P\rangle$$

$$g'''(t)=\langle D^3f_{P+t(X-P)}\cdot(X-P)\cdot(X-P),X-P\rangle$$

Nuevamente, consideramos la fórmula de Taylor en una variable para la función g:

$$g(1)=g(0)+g'(0)+\frac{1}{2}g''(0)+\frac{1}{6}g'''(c)$$

con c algún punto en el intervalo $(0,1)$. Entonces

$$f(X) = f(P) + \langle \nabla f_P, X-P\rangle + \frac{1}{2}\langle Hf_P(X-P), X-P\rangle + \frac{1}{6}\langle D^3 f_{P+c(X-P)} \cdot (X-P)\cdot(X-P), X-P\rangle$$

donde $c \in (0,1)$ con lo cual $C = P + c(X-P)$ está efectivamente en el segmento que une P con X. Observemos además que:

$$|\langle D^3 f_C(X-P)\cdot(X-P), X-P\rangle| \leq ||D^3 f_C(X-P)||_\infty ||X-P||^2$$

y, por el lema previo,

$$||D^3 f_C(X-P)||_\infty \leq \sum_{i=1}^{n}\left(\left|\left|H\frac{\partial f}{\partial x_i}|_C\right|\right|_\infty^2\right)^{\frac{1}{2}}$$

Notemos que $H\frac{\partial f}{\partial x_i}$ involucra, para cada i, las derivadas de orden 3 de f. Como cada una de ellas es una función continua por hipótesis, para un entorno de P, es también, una función acotada. Entonces, si X está cerca de P, vale que

$$\left|\left|H\frac{\partial f}{\partial x_i}|_C\right|\right|_\infty \leq n\cdot \max\left|\frac{\partial^3 f}{\partial x_j\ \partial x_k\ \partial x_l}(C)\right| \leq M$$

Luego, se tiene que

$$|R_P(X-P)| = \frac{1}{6}|\langle D^3 f_C(X-P)\cdot(X-P), X-P\rangle| \leq M||X-P||^3$$

de donde se deduce inmediatamente el teorema, pues

$$\frac{|R_P(X-P)|}{||X-P||^2} \leq M||X-P||$$

para cualquier X lo suficientemente cerca de P. □

Teorema 34 (Multiplicadores de Lagrange en $\mathbb{R}^3$). *Sea $f : \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable y $g : \mathbb{R}^3 \to \mathbb{R}$ una función C^1. Sea $P = (p_1, p_2, p_3)$ tal que $\nabla_g(P) \neq (0,0,0)$. Si P es un extremo de f restringida a la superficie de nivel $S = \{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = g(P)\}$, entonces*

$$\nabla_f(P) = \lambda \nabla_g(P)$$

O, equivalentemente, $\nabla_f(P)$ y $\nabla_g(P)$ son linealmente dependientes.

Demostración. Sin pérdida de generalidad, supongamos que la coordenada de g que no se anula es la última. O sea $\frac{\partial g}{\partial z} \neq 0$. Por esto y porque g es C^1, podemos aplicarle a g el Teorema de la función Implícita, que nos permite afirmar que existen una función $\varphi : \mathbb{R}^2 \to \mathbb{R}$ biyectiva y derivable y dos entornos U y V centrados en (p_1, p_2) y p_3 respectivamente tales que $\varphi(p_1, p_2) = p_3$.

Consideremos ahora $G(x,y) = g(x, y, \varphi(x,y))$. En un entorno del punto (p_1, p_2), G es una función constante pues $(p_1, p_2, \varphi(p_1, p_2)) \in S$. Consecuentemente, su gradiente en (p_1, p_2) se anula: $\nabla_G(p_1, p_2) = 0$. Ahora, mediante la regla de la cadena, expandamos la expresión de sus derivadas parciales.

$$G_x(P) = g_x(P)\cdot 1 + g_y(P)\cdot 0 + g_z(P)\cdot \varphi'(p_1, p_2) = 0$$

$$G_x(P) = \langle \nabla_g(P), (1, 0, \varphi'(p_1, p_2))\rangle = 0 \tag{2}$$

Operando análogamente con su derivada parcial con respecto a y,

$$G_y(P) = g_x(P)\cdot 0 + g_y(P)\cdot 1 + g_z(P)\cdot \varphi'(p_1, p_2) = 0$$

$$G_y(P) = \langle \nabla_g(P), (0, 1, \varphi'(p_1, p_2)) \rangle = 0 \tag{3}$$

Sea Π el plano generado por los vectores $(1, 0, \varphi'(p_1, p_2))$ y $(0, 1, \varphi'(p_1, p_2))$. Las ecuaciones (2) y (3) nos muestran que el vector gradiente de la función g evaluado en el punto P es ortogonal al plano Π.

Dejando esto de lado momentáneamente, consideremos ahora la función $F(x, y) = f(x, y, \varphi(x, y))$. Como f restringida a S tiene un extremo en P (supongamos que es un máximo. El caso de mínimo es análogo), vale que

$$F(p_1, p_2) = f(p_1, p_2, \varphi(p_1, p_2)) = f(p_1, p_2, \varphi(p_1, p_2)) \geq f(x_1, x_2, x_3)$$

para cualquier terna (x_1, x_2, x_3) suficientemente cercana a (p_1, p_2, p_3). Pero como vale para cualquier terna, en particular vale para las ternas de la forma $(x_1, x_2, \varphi(x_1, x_2))$. O sea, vale que

$$F(p_1, p_2) = f(p_1, p_2, \varphi(p_1, p_2)) \geq f(x_1, x_2, \varphi(x_1, x_2)) = F(x_0, x_1)$$

para cualquier par (x_0, x_1) suficientemente cercano a (p_1, p_2). Pero esto significa que la función F restringida a S tiene un extremo local en S, de lo que se deduce que su vector gradiente en (p_1, p_2) se anula: $\nabla_F(p_1, p_2) = 0$. Aplicando la regla de la cadena, obtenemos que:

$$F_x(P) = f_x(P) \cdot 1 + f_y(P) \cdot 0 + f_z(P) \cdot \varphi'(p_1, p_2) = 0$$
$$F_x(P) = \langle \nabla_f(P), (1, 0, \varphi'(p_1, p_2)) \rangle = 0 \tag{4}$$

Operando análogamente con su derivada parcial con respecto a y,

$$F_y(P) = f_x(P) \cdot 0 + f_y(P) \cdot 1 + f_z(P) \cdot \varphi'(p_1, p_2) = 0$$
$$F_y(P) = \langle \nabla_f(P), (0, 1, \varphi'(p_1, p_2)) \rangle = 0 \tag{5}$$

Al igual que con G, las ecuaciones (4) y (5) nos muestran que el vector gradiente de la función f evaluado en el punto P es ortogonal al mismo plano Π al que ya se demostró que era ortogonal el gradiente de g. Dado que $\nabla_f(P)$ y $\nabla_g(P)$ son ortogonales al mismo plano, deben ser paralelos, o, equivalentemente, linealmente independientes:

$$\nabla_f(P) = \lambda \nabla_g(P)$$

$\square$

Teorema 35 (Multiplicadores de Lagrange)**.** *Para hallar los puntos críticos de una función $f : \mathbb{R}^n \to \mathbb{R}$ (f diferenciable) restringida a la superficie de nivel $g(X) = c$ dada por una función $g : \mathbb{R}^n \to \mathbb{R}$ (con $g \in C^1$) basta estudiar los puntos críticos de la función de $n + 1$ variables dada por:*

$$f(X) - \lambda(g(X) - c)$$

Más precisamente, si $P \in \mathbb{R}^n$ es un extremo de f restringida a la superficie de nivel de $g(X) = c$, y $\nabla g_P \neq \mathbb{O}$, entonces existe $\lambda \in \mathbb{R}$ tal que

$$\nabla f_P = \lambda \nabla g_P$$

Demostración. Supongamos que $\nabla g_P \neq \mathbb{O}$. Entonces, por el teorema de la función implícita, existen una bola $B \subseteq \mathbb{R}^{n-1}$ y una parametrización $\varphi : B \to \mathbb{R}$ de la superficie S en un entorno de P. Llamamos $Z_0 \in \mathbb{R}^{n-1}$ al centro de la bola B. Para simplificar, supongamos que la derivada de g que no se anula es la última, de manera de que $(Z_0, \varphi(Z_0)) = P$, y además $g(Z, \varphi(Z)) = c$ para todo $Z \in B$.
Ahora si P es un máximo de f restringido a S, debe valer que $f(P) \geq f(X) \forall X \in S$ en un entorno de P. O sea,

$$f(Z_0, \varphi(Z_0)) \geq f(Z, \varphi(Z))$$

para todo Z suficientemente cerca de Z_0. Entonces, la función $h(Z) = f(Z, \varphi(Z))$ tiene un extremo local en Z_0, con lo cual su gradiente se anula en el punto o, equivalentemente, todas sus derivadas parciales son nulas en el punto. Pero, por regla de la cadena

$$\frac{\partial h(Z)}{\partial z_i}|_{Z_0} = \frac{\partial}{\partial z_i} f(Z, \varphi(Z))_{Z_0} = \langle \nabla f_P, (E_i, \frac{\partial \varphi}{\partial z_i}(Z_0)) \rangle$$

Esto significa que el gradiente de f en P es perpendicular a todos los vectores generadores del plano tangente a la superficie, y, en consecuencia, tiene que ser paralelo a la normal. Como la normal es el gradiente de g en el punto, se tiene

$$\nabla f_P = \lambda \nabla g_P$$

□

8 Integrales en $\mathbb{R}$

Teorema 36 (Método de integración por partes)**.**

$$\int (u \cdot v') = u \cdot v - \int (u' \cdot v)$$

Teorema 37. *Si $f : [a, b] \to \mathbb{R}$ es acotada, entonces f es integrable en $[a, b]$ si y sólo si para cualquier $\varepsilon > 0$ existe una partición de P de $[a, b]$ tal que*

$$S(f, P) - I(f, P) < \varepsilon$$

Demostración.
$\Leftarrow$) Supongamos que vale que para todo ε, existe P partición tal que $S(f, P) - I(f, P) < \varepsilon$. Probar que la función es integrable en el $[a, b]$ equivale, por definición, a probar que $\sup(\{I(f, Q)\}) = \inf(\{S(f, Q')\})$ para todo par Q, Q' particiones de $[a, b]$.

- $\sup(\{I(f, Q)\}) \leq \inf(\{S(f, Q')\})$ es trivial pues consideremos Q'' partición de $[a, b]$ que refina tanto a Q como a Q'. Entonces, vale que

  $$I(f, Q) \leq I(f, Q'') \leq S(f, Q'') \leq S(f, Q')$$

  $$I(f, Q) \leq S(f, Q')$$

  Como **cualquier** suma inferior es menor o igual que **cualquier** suma superior, en particular la mayor de las sumas inferiores es menor que la menor de las sumas superiores.

- $\sup(\{I(f, Q)\}) \geq \inf(\{S(f, Q')\})$. Por hipótesis, sabemos existe P partición de $[a, b]$ tal que que $S(f, P) - I(f, P) < \varepsilon$. Además, dado que $\sup(\{I(f, Q)\}) \leq I(f, P)$ y que $\inf(\{S(f, Q')\}) \leq S(f, P)$, obtenemos que

  $$\sup(I(f, Q)) - \inf(S(f, Q')) \leq S(f, P) - I(f, P) < \varepsilon$$

  Con lo cual $\sup(I(f, Q)) \geq \inf(S(f, Q'))$

Consecuentemente, $\sup(I(f, Q)) = \inf(S(f, Q'))$ y f es integrable.

$\Rightarrow$) Supongamos ahora que f es integrable. Por definición de integrabilidad, esto es que existe un valor tal que

$$\int_a^b f = I_*(f) = I^*(f)$$

siendo $I_*(f)$ la integral inferior de f e $I^*(f)$ su integral superior.
Nuevamente, aplicando la definición de integrales superior e inferior, obtenemos que, por ser f integrable,

$$\int_a^b f = \sup\{I(f, P) : P \text{ partición de } [a, b]\} = \inf\{S(f, P) : P \text{ partición de } [a, b]\}$$

Por las propiedades de ínfimo y supremo, dado $\varepsilon > 0$ existen particiones P_1 y P_2 tales que

- $\displaystyle\int_a^b f(x)\,\mathrm{d}x - I(f, P_1) < \frac{\varepsilon}{2}$
- $\displaystyle S(f, P_2) - \int_a^b f(x)\,\mathrm{d}x < \frac{\varepsilon}{2}$

Sumando miembro a miembro las anteriores desigualdades, obtenemos

$$S(f, P_2) - I(f, P_1) < \varepsilon$$

Consideremos ahora la partición $P = P_1 \cup P_2$. Como P refina a P_1, $I(f, P_1) < I(f, P)$. Similarmente, como P refina a P_2, $S(f, P_2) > S(f, P)$. Consecuente,

$$S(f, P) - I(f, P) < S(f, P_2) - I(f, P_1) < \varepsilon$$

Por la propiedad transitiva de la desigualdad estricta se obtiene la tesis del teorema. □

Proposición 38. *Sean $f, g : [a, b] \to \mathbb{R}$ dos funciones integrables en $[a, b]$ tales que para todo $x \in [a, b]$, $g(x) \leq f(x)$. Entonces*

$$\int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x$$

Proposición 39. *$f : [a, b] \to \mathbb{R}$ es acotada e integrable entonces*

$$\left| \int_a^b f(x)\,\mathrm{d}x \right| \leq \int_a^b |f(x)|\,\mathrm{d}x$$

Demostración. Por definición de módulo, vale que:

- Si $|\int_a^b f(x)\,\mathrm{d}x| = \int_a^b f(x)\,\mathrm{d}x$, basta observar que para todo x, $f(x) \leq |f(x)|$. Consecuentemente, $|\int_a^b f(x)\,\mathrm{d}x| = \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x$ (por la proposición anterior).
- Si $|\int_a^b f(x)\,\mathrm{d}x| = -\int_a^b f(x)\,\mathrm{d}x = \int_a^b -f(x)\,\mathrm{d}x$, basta observar que para todo x, $-f(x) \leq |-f(x)| \leq |f(x)|$. Consecuentemente $|\int_a^b f(x)\,\mathrm{d}x| = \int_a^b -f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x$ (por la proposición anterior).

□

Teorema 40 ((Continua $\circ$ Integrable) $\Rightarrow$ Integrable)**.** *Sea $h : [a, b] \to [c, d]$ integrable y $\varphi : [c, d] \to \mathbb{R}$ continua. Entonces $\varphi \circ h : [a, b] \to \mathbb{R}$ es integrable.*

Demostración. Nomenclatura para la demostración:

- P es una partición particular de $[a, b]$.
- Δ_i es el i-ésimo intervalo de la partición P.
- M_i es el supremo de h en Δ_i.
- m_i es el ínfimo de h en Δ_i.
- M_i^* es el supremo de $\varphi \circ h$ en Δ_i.
- m_i^* es el ínfimo de $\varphi \circ h$ en Δ_i.

En base a la nomenclatura definida hasta ahora, observemos que

$$M_i - m_i = \sup\{h(x) : x \in \Delta_i\} - \inf\{h(y) : y \in \Delta_i\} = sum\{h(x) - h(y) : x, y \in \Delta_i\}$$

Esto es porque en cada intervalo Δ_i la diferencia entre el valor más grande y el más chico de h coincide con la diferencia más grande de valores de h. Análogamente,

$$M_i^* - m_i^* = \sup\{\varphi(h(x)) - \varphi(h(y)) : x, y \in \Delta_i\}$$

Como h es integrable, existe P una partición de $[a,b]$ tal que

$$S(h,P) - I(h,P) = \sum (M_i - m_i)\Delta_i \leq \varepsilon_1 \tag{6}$$

A su vez, como φ es uniformemente continua, dado $\varepsilon_2 > 0$, existe $\delta > 0$ tal que para cualquier par de valores $t, s \in [c,d]$, $|t-s| < \delta$ implica que $|\varphi(t) - \varphi(s)| < \varepsilon_2$.

Lo que nosotros queremos probar, concretamente es que la siguiente cantidad es arbitrariamente pequeña

$$S(\varphi \circ h, P) - I(\varphi \circ h, P) = \sum (M_i^* - m_i^*)\Delta_i$$

Para esto, observemos cada uno de los intervalos de la partición. Para un índice i dado, existen dos posibilidades:

- $M_i - m_i < \delta$
- $M_i - m_i \geq \delta$

Definimos entonces al conjunto de índices A como el conjunto que contiene a todos los índices que caen en el primer caso y al conjunto B como el conjunto de índices que caen en el segundo caso. Formalizando

- $A = \{i : M_i - m_i < \delta\}$
- $B = \{i : M_i - m_i \geq \delta\}$

Retomando lo que queremos probar que es demostrar que es arbitrariamente pequeño,

$$\sum (M_i^* - m_i^*)\Delta_i = \sum_{i \in A} (M_i^* - m_i^*)\Delta_i + \sum_{i \in B} (M_i^* - m_i^*)\Delta_i \tag{7}$$

Analicemos el primero de los términos. Para todo $i \in A$ (o sea, si i cayera en el primer caso), entonces $M_i^* - m_i^* < \delta$ por la continuidad uniforme de φ. Consecuentemente,

$$\sum_{i \in A} (M_i^* - m_i^*)\Delta_i < \varepsilon_2 \sum_{i \in A} \Delta_i \leq \varepsilon_2 (b-a)$$

Es decir, eligiendo ε_2 lo suficientemente pequeño, podemos asegurar que las sumas superiores e inferiores están tan cerca como queramos en todos los índices del conjunto A.

Analicemos ahora el segundo de los términos. Para todo $i \in B$ (o sea, si i cayera en el segundo caso), como φ es una función continua en un compacto, alcanza su máximo. O sea, existe $M > 0$ tal que $\varphi(x) \leq M$ para todo $x \in [c,d]$. Como vale para todo el intervalo, vale para cualquier subconjunto que esté contenido en $[c,d]$, en particular el intervalo de la partición Δ_i. Esto nos permite afirmar que

$$M_i^* - m_i^* = \sup\{\varphi(h(x)) - \varphi(h(y)) : x, y \in \Delta_i\} \leq 2M$$

Luego, podemos acotar de la siguiente forma:

$$\sum_{i \in B} (M_i^* - m_i^*)\Delta_i \;\leq\; 2M \sum_{i \in B} \Delta_i \;\leq\; \frac{2M}{\delta} \sum_{i \in B} \delta \Delta_i \;\leq\; \frac{2M}{\delta} \sum_{i \in B} (M_i - m_i)\Delta_i \;<\; \frac{2M}{\delta}\varepsilon_1$$

por la ecuación (6). Es decir, si elegimos ε_1 lo suficientemente pequeño (o sea, elegimos la partición P lo suficientemente fina), podemos lograr que la diferencia entre las sumas superiores y las sumas inferiores estén tan cerca como queramos en todos los índices del conjunto B.

Retomando la ecuación (7)

$$\sum_{i \in A} (M_i^* - m_i^*)\Delta_i + \sum_{i \in B} (M_i^* - m_i^*)\Delta_i \leq \varepsilon_2 (b-a) + \frac{2M}{\delta}\varepsilon_1 < \varepsilon$$

Esto demuestra que las sumas superiores están tan cerca de las sumas inferiores como queramos, lo que prueba que $\varphi \circ h$ es integrable. $\square$

Proposición 41. *Sea $f : [a, b] \to \mathbb{R}$ continua. Entonces f es integrable.*

Demostración. La demostración de esta proposición se basa en el teorema anterior considerando:

- $h(x) = x$ que, trivialmente, es integrable.
- $\varphi(x) = f(x)$ que es continua por hipótesis.

Entonces, por el teorema anterior, $f(x) = \varphi(h(x)) = (\varphi \circ h)(x)$ es integrable. $\square$

Proposición 42. *Sean $f, g : [a, b] \to \mathbb{R}$ integrables. Entonces*

- *$f + g$ es integrable.*
- *f^n es integrable.*
- *$f \cdot g$ es integrable. (OJO. No necesariamente vale que $\int f \cdot g = \int f \cdot \int g$)*

Teorema 43 (TFC)**.** *Sea $f : [a, b] \to \mathbb{R}$ una función continua. Dado $x \in [a, b]$, sea $F : [a, b] \to \mathbb{R}$ definida como*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

*Entonces F (llamada **función primitiva** de f) es continua en $[a, b]$, derivable en (a, b) y para todo $x \in (a, b)$ se tiene que*

$$F'(x) = f(x)$$

Demostración. En primer lugar, es necesario probar que F es continua en los bordes del intervalo (o sea, en los puntos a y b). Para eso, es necesario probar que

$$\lim_{x \to a^+} F(x) = F(a) \qquad\qquad \lim_{x \to b^-} F(x) = F(b)$$

O, equivalentemente,

$$\lim_{x \to a^+} |F(x) - F(a)| = 0 \qquad\qquad \lim_{x \to b^-} |F(x) - F(b)| = 0$$

Comencemos por demostrar esto. Sea $s = \max\{|f(t)| : t \in [a, b]\}$ (que es finito por ser f continua). Consideremos entonces P es una partición del intervalo $[a, x]$ y M_i como el supremo de $|f(t)|$ en el intervalo Δ_i. Vale entonces que

$$\begin{aligned} |F(x) - F(a)| = |\int_a^x f - \int_a^a f| = |\int_a^x f| \leq \int_a^x |f| \leq \inf S(|f|, P) \\ \leq S(|f|, P) = \sum_i M_i \Delta_i \leq s \sum_i \Delta_i = s(x - a) \end{aligned}$$

De este modo se demuestra que $F(x) \to F(a)$ cuando $x \to a^+$. De forma análoga, consideremos

$$|F(b) - F(x)| = |\int_a^b f - \int_a^x f| = |\int_x^b f| \leq \int_x^b |f| = \inf S(|f|, P) \leq s(b - x)$$

que demuestra que $F(x) \to F(b)$ cuando $x \to b^-$.

Ahora queremos demostrar que $F'(x) = f(x)$. Entonces, por definición, esto equivale a demostrar que

$$\lim_{h \to 0} \frac{F(x + h) - F(x)}{h} - f(x) = 0$$

Trabajemos expandamos la definición de F y simplifiquemos algunos términos

$$\begin{aligned}
F'(x) - f(x) &= \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} - f(x) \\
&= \lim_{h\to 0} \frac{\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t}{h} - f(x) \\
&= \lim_{h\to 0} \frac{\cancel{\int_a^x f(t)\,\mathrm{d}t} + \int_x^{x+h} f(t)\,\mathrm{d}t - \cancel{\int_a^x f(t)\,\mathrm{d}t}}{h} - f(x) \\
&= \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t}{h} - f(x)
\end{aligned}$$

Consideremos ahora la siguiente escritura mágica para $f(x)$:

$$f(x) = f(x)\cdot 1 = f(x)\frac{h}{h} = \frac{f(x)\int_x^{x+h} 1}{h} = \frac{\int_x^{x+h} f(x)\,\mathrm{d}x}{h}$$

Retomando la expansión con la que veníamos trabajando,

$$\begin{aligned}
&= \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t}{h} - f(x)r \\
&= \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t}{h} - \frac{\int_x^{x+h} f(x)\,\mathrm{d}x}{h} \\
&= \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t - \int_x^{x+h} f(x)\,\mathrm{d}x}{h} \\
&= \lim_{h\to 0} \frac{\int_x^{x+h} f(t) - f(x)\,\mathrm{d}t}{h}
\end{aligned}$$

Intentemos ahora resolver este límite por definición: la variable de integración t, al estar siendo integrada entre x y $x+h$ verifica que $x \le t \le x+h]$. Entonces $|t-x| < |x+h-x| = |h| < \delta$. Por definición de límite y continuidad de f, como sabemos que x dista de t en menos que delta, podemos afirmar que existe $\varepsilon > 0$ tal que $|f(t) - f(x)| < \varepsilon$.

Entonces, suponiendo $h > 0$ y acotando apropiadamente,

$$\left| \frac{\int_x^{x+h} f(t) - f(x)\,\mathrm{d}t}{h} - 0 \right| \tag{8}$$

$$\le \frac{\int_x^{x+h} |f(t) - f(x)|\,\mathrm{d}t}{|h|} \tag{9}$$

$$\le \frac{\int_x^{x+h} \varepsilon}{h} \tag{10}$$

$$= \frac{\varepsilon(\cancel{x} + h - \cancel{x})}{h} \tag{11}$$

$$= \frac{\varepsilon \cancel{h}}{\cancel{h}} \tag{12}$$

$$= \varepsilon \tag{13}$$

El caso $h < 0$ es análogo: el denominador es negativo, pero al darse vuelta la integral del numerador, aparece un signo menos arriba y se cancelan. □

Demostración Alternativa. Se trabaja igual que la demostración anterior hasta llegar a que

$$F'(x) = \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t}{h}$$

Dado que f es una función continua (por hipótesis) y el intervalo $[x, x+h]$ es un compacto, la función alcanza sus valores máximos y mínimos (llamémoslos M_h y m_h). Sin embargo, cuando $h \to 0$, el intervalo $[x, x+h]$ tiende al punto x y, consecuentemente, tanto m_h como M_h tienden a $f(x)$. Entonces

$$\lim_{h\to 0}\left(m_h \leq \frac{\int_x^{x+h} f}{h} \leq M_h\right) = (f(x) \leq F'(x) \leq f(x))$$

Consecuentemente $F'(x) = f(x)$ □

Teorema 44 (Regla de Barrow)**.** *Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua y sea $G : [a,b] to \mathbb{R}$ continua (¿¿¿Posta continua??? Para mi G no tiene que ser continua) tal que $G'(x) = f(x)$ para todo $x \in [a,b]$. Entonces*

$$G(b) - G(a) = \int_a^b f(x)\,\mathrm{d}x$$

Demostración. Como f es una función continua, por el Teorema Fundamental del Cálculo Integral, debe existir $F(x) = \int_a^x f$ tal que $F'(x) = f$. Pero, como $F'(x) = f(x) = G'(x)$, necesariamente $G(x) = F(x) + c$ para algún $c \in \mathbb{R}$ (como consecuencia del teorema de unicidad de la derivada + constante). Entonces

$$G(b) - G(a) = F(b) + \cancel{c} - F(a) - \cancel{c} = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x - \int_a^a f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x - 0 = \int_a^b f(x)\,\mathrm{d}x$$

Por la propiedad transitiva de la relación de equivalencia de igualdad, considerando sólo los extremos de la cadena de igualdades obtenemos la tesis deseada. □

Teorema 45 (Teorema de Cambio de Variable)**.** *Sea $\varphi : [\alpha, \beta] \to [a,b]$ una función C^1 tal que $\varphi(\alpha) = a$ y $\varphi(\beta) = b$. Además, sea $f : [a,b] \to \mathbb{R}$ continua en $[a,b]$. Entonces $(f \circ \varphi) \cdot \varphi' : [\alpha, \beta] \to \mathbb{R}$ es integrable en $[\alpha, \beta]$ y*

$$\int_a^b f(t)\,\mathrm{d}t = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(t)\,\mathrm{d}t = \int_\alpha^\beta (f \circ \varphi)(u) \cdot \varphi'(u)\,\mathrm{d}u$$

Demostración. Como f es integrable, se le puede aplicar el Teorema Fundamental del Cálculo Integral. Consecuentemente, existe $F = \int_a^x f$ tal que $F'(x) = f(x)$. Consideremos ahora la derivada de la función $(F \circ \varphi)$. Por regla de la cadena, vale que

$$(F \circ \varphi)' = (F' \circ \varphi) \cdot \varphi' = (f \circ \varphi) \cdot \varphi' \tag{14}$$

Por otro lado, aplicando la Regla de Barrow a f, sabemos que

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a) = F(\varphi(\beta)) - F(\varphi(\alpha)) = (F \circ \varphi)(\beta) - (F \circ \varphi)(\alpha)$$

Sin embargo, por lo visto en la ecuación (14), sabemos que $(F \circ \varphi)$ es una primitiva de $(f \circ \varphi) \cdot \varphi'$. Nuevamente, aplicando la Regla de Barrow,

$$\int_a^b f(t)\,\mathrm{d}t = \ldots = (F \circ \varphi)(\beta) - (F \circ \varphi)(\alpha) = \int_\alpha^\beta (f \circ \varphi)(u) \cdot \varphi'(u)\,\mathrm{d}u$$

□

Proposición 46 (Criterios de Comparación)**.** *Sean $f, g : [a,b) \to \mathbb{R}$ funciones integrables (b puede ser $+\infty$) y no negativas.*

1. *Si $f(t) \leq g(t)$ para todo $t \in [a,b)$, entonces si $\int_a^b g$ converge, $\int_a^b f$ también converge.*

2. *Si $f(t) \geq g(t)$ para todo $t \in [a,b)$, entonces si $\int_a^b f$ diverge, $\int_a^b g$ también diverge.*

Demostración.
1. Consideremos $F = \int_a^x f$ y $G = \int_a^x g$. Por el Teorema Fundamental del Cálculo Integral, ambas son crecientes en $[a, b)$ (por ser $f, g \geq 0$). Como además, por hipótesis, $f \leq g$, vale que $F(x) \leq G(x)$ para todo $x \in [a, b)$. Luego, como $\int_a^b g$ converge, vale que $G(x) = \int_a^x g \leq \int_a^b g$. Consecuentemente,

$$F(x) \leq G(x) \leq \int_a^b x$$

Por transitividad de la desigualdad,

$$F(x) \leq \int_a^b g$$

Esto prueba que F está acotada y, como era creciente, su límite $x \to b+$ existe y es finito.

2. Como $f(x) \leq g(x)$ para todo $x \in [a, b)$, entonces $\int_a^b f \leq \int_a^b g$. Pero, $\int_a^b f = \infty$, entonces por transitividad del operador $\leq$ tenemos que,

$$\infty = \int_a^b f \leq \int_a^b g \Rightarrow \int_a^b g = \infty$$

$\square$

9 Integrales Múltiples

Teorema 47 (Fubini simplificado)**.** *Sea* $R = [a, b]x[c, d]$ *un rectángulo en* $\mathbb{R}^2$*, y* $f : R \to \mathbb{R}$ *una función integrable en* R*.*

- 1. *Si para todo* $x \in [a, b]$*,* $f_x : [c, d] \to \mathbb{R}$ *es integrable, entonces*

$$\int\int_R f = \int_a^b \int_c^d f(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

- 2. *Si para todo* $y \in [c, d]$*,* $f_y : [a, b] \to \mathbb{R}$ *es integrable, entonces*

$$\int\int_R f = \int_c^d \int_a^b f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

Demostración. Nomenclatura para la demostración:

- $\Delta_i = [x_i, x_{i+1}]$ es una partición del intervalo $[a, b]$.
- $\Delta'_j = [y_j, y_{j+1}]$ es una partición del intervalo $[c, d]$.
- $m_{ij}(f)$ es el ínfimo de f en el rectángulo R_{ij}.
- $m_j(f_x)$ es el ínfimo de f_x en el intervalo Δ'_j
- $\mathbb{I} : [a, b] \to \mathbb{R}$ es la integral de f_x con respecto de y.
- $I_*(f)$ es el supremo de las sumas inferiores de f (llamado integral inferior de f).

Supongamos que f_x es integrable para todo $x \in [a, b]$. Entonces, acorde a la nomenclatura definida arriba,

$$\mathbb{I}(x) = \int_c^d f_x(y)\,\mathrm{d}y = \int_c^d f(x, y)\,\mathrm{d}y$$

Si $x \in R_{ij}$, entonces $m_j(f_x)$ es mayor o igual que el ínfimo de f en R_{ij} pues $m_j(f_x) = \inf\{f(x, y) : x$ fijo $, y \in \Delta'_j\}$ y el ínfimo en un subconjunto (en este caso $\{x\}$ x Δ'_j) siempre es mayor o igual que el ínfimo en el conjunto original (en este caso Δ_i x Δ'_j). ‘’ Se deduce así que, para todo $x \in \Delta_i$

$$\sum_j m_{ij}(f)\Delta'_j \leq \sum_j m_j(f_x)\Delta'_j$$

Por definición, $\sum_j m_j(f_x)\Delta'_j$ es la suma inferior de f_x. Y dado que supusimos que f_x es integrable, entonces su integral coincide con su integral inferior (por definición de integral de Riemann). De este modo, podemos continuar la desigualdad anterior de la siguiente forma:

$$\sum_j m_{ij}(f)\Delta'_j \leq \sum_j m_j(f_x)\Delta'_j \leq I_*(f_x) = \int_c^d f_x(y)\,\mathrm{d}y = \mathbb{I}(x) \tag{15}$$

El resultado obtenido en (15) nos permite afirmar que

$$I(f, P) = \sum_{i,j} m_{ij}(f)\Delta_i\Delta'_j = \sum_i \left(\sum_j m_{ij}(f)\Delta'_j\right)\Delta_i \leq \sum_i m_i(\mathbb{I})\Delta_i \leq I_*(\mathbb{I})$$

De lo que se obtiene $I(f,P) \leq I_*(\mathbb{I})$. De forma análoga se demuestra que $I^*(\mathbb{I}) \leq S(f,P)$. Y como siempre vale que $I_* \leq I^*$, podemos afirmar que

$$I(f,P) \leq I_*(\mathbb{I}) \leq I^*(\mathbb{I}) \leq S(f,P)$$

Como f es integrable, los extremos se pueden hacer tan próximos como queramos refinando la partición P y, en consecuencia, $\mathbb{I}(x)$ es integrable. Además,

$$\int\int_R f = I_*(\mathbb{I}) = I^*(\mathbb{I}) = \int_a^b \mathbb{I} = \int_a^b \int_c^d f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$\square$

Teorema 48 (Teorema del Cambio de Variable)**.** *Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ inyectiva de clase C^1 (o sea, $J_T \neq 0$). Sea $D^* \subseteq \mathbb{R}^2$ acotado y $f : D = T(D^*) \to \mathbb{R}$ integrable. Si D_T es inversible en D^* y $J_T = |det(D_T)|$, entonces:*

$$\int_D f = \int_{D^*} (f \circ T) \cdot J_T$$